



# Deploying A Functional caGrid Node

## caBIG™ Deployment Teleconference

**August 7, 2008**

**George Komatsoulis, Ph.D.**

Deputy Director

National Cancer Institute Center for Biomedical  
Informatics and Information Technology



# Agenda



## Deploying a Functional Grid Node

- What is a “functional” caGrid node?
- Why deploy a functional caGrid node?
- How can you successfully deploy a functional caGrid node?
- When do you have to deploy a functional caGrid node?
- What resources and support are available to you?

# What is a “functional” caGrid node?



- **A “functional” caGrid node is ...**  
...a node that you establish to securely connect your Center’s select data or analytic services to a network of researchers and organizations who are also sharing *their* data and analytic services.

# Why deploy a functional caGrid node?



- A functional caGrid node will enable you to:
  1. Establish an interoperable resource infrastructure at your Center to facilitate new avenues of data sharing and exchange
  2. Securely share protected clinical/scientific data with a network of collaborators
  3. Join in fostering a connected research community by gaining and providing access to data or analysis services

# How can you successfully deploy a functional caGrid node?



- Three ways to deploy a functional Grid node include:
  1. Install the Life Sciences Distribution (LSD) Bundle
    - Analogous to “Standard Install”
  2. Install any caBIG™ application
    - Analogous to “Custom Install”
  3. Adapt one of your existing applications
    - Analogous to building your own installer
    - Software engineering approach

# Path 1: Install the LSD Bundle



## The Steps Involved

1. Install the LSD bundle
2. Use caBIG™ supplied tool kit to register services on the Grid using either your own data or data or the reference data included in the applications

## Key Features of the Approach

- Simplest approach technically
- Well tested approach that requires the least effort and configuration
- Installed applications include:
  - caArray (gene array data management repository)
  - caGWAS (genome wide association repository)
  - caTissue (biospecimen repository)
  - National Cancer Imaging Archive (imaging tool)
  - CTODS (clinical trials data repository)

## Support Available Through caBIG™

- Documentation: Installation guide, End user documentation
- caGrid deployment team
- Webinars offered through the Deployment Advisory Center and available on caBIG™ web site

## Path 2: Install any caBIG™ application



### The Steps Involved

1. Install a Grid-compatible version of any caBIG™ application
2. Use caBIG™ supplied tool kit to register services on the Grid using either your own data or the reference data included in the applications

### Key Features of the Approach

- Not necessarily as technically simple as Path 1
- Choose the application to be deployed and make decisions about the infrastructure to support it based on recommendations available through caBIG™

### Support Available Through caBIG™

- Documentation: Installation guide, End user documentation
- caGrid Deployment Team
- caBIG™ Knowledge Centers
- caBIG™ Support Service Providers

# Path 3: Adapt one of your existing applications



## The Steps Involved

1. Adapt one of your existing applications
2. Follow the process for creating a Grid Service
3. Share actual data or test data

## Key Features of the Approach

- Technically least simple
- Requires software development/ engineering expertise

## Support Available Through caBIG™

- caBIG™ Software Development Tools (caCore, SDK, and Introduce)
- Adaptation Design Patterns
- caBIG™ Cross-cutting Workspaces (Architecture and Vocabulary & Common Data Elements)
- caBIG™ Knowledge Centers
- caBIG™ Support Service Providers
- caBIG™ Program staff will work with you individually to support adaptation of your chosen application



# The Paths: General Considerations



- Irrespective of the path you choose:
  - Your caGrid node deployment will be a critical first step in being able to take advantage of caGrid.
  - Your node on the Grid will join a network of caGrid nodes that together will form the largest research infrastructure ever created for data access and exchange.

# When do you have to deploy a functional caGrid node?



- By September 30, 2008: Deploy a functional Grid node and share either reference or actual data or analytic service on the Grid
- Post-September 30, 2008: Share actual data via the Grid node based on timelines provided by each Center in the Center caBIG™ Goals and Implementation Plan

# What resources and support are available to you?



1. Demonstrations of caGrid node deployment webcast (will be made available on caGrid page on caBIG™ web site)
2. caGrid Deployment Team (Contact info: [caBIGconnect@cancer.gov](mailto:caBIGconnect@cancer.gov))
3. caGrid User Group Listserv (CAGRID\_USERS-L@LIST.NIH.GOV)
4. caGrid Wiki (<http://www.cagrid.org>)
5. caBIG™ Enterprise Support Network (<https://cabig.nci.nih.gov/esn>)
  - a. Knowledge Centers
    - caGrid, Clinical Trials Management Systems, Molecular Analysis Tools, Tissue/Biospecimen Banking and Technology Tools, Data Sharing and Intellectual Capital, Vocabulary
    - The URL to the Knowledge Center web site will be made available as soon as site setup is complete.
  - b. caBIG™ Support Service Providers